<110> Merja PENTTILA et al.

<120> PROCESS FOR PARTITIONING OF PROTEINS

<130> 0933-0170P

<140> US 09/936,823
<141> 2001-10-24

<150> PCT/FI00/00249
<151> 2000-03-24

<150> FI 19991782
<151> 1999-08-20

<150> FI 19990667
<151> 1999-03-25

<160> 46

<170> PatentIn Ver. 2.2

<210> 1
<211> 428
<212> DNA
<213> Trichoderma reesei

<220>
<221> intron
<222> (167)..(236)

<220>
<221> intron
<222> (323)..(386)

<220>
<221> misc_feature
<223> Coding sequence of hfb1

<400> 1
atgaagttct tcgccatcgc cgctctcttt gccgccgctg ccgttgccca gcctctcgag 60
gaccgcagca acggcaacgg caatgtttgc cctcccggcc tcttcagcaa cccccagtgc 120
tgtgccaccc aagtccttgg cctcatcggc cttgactgca aagtccgtaa gttgagccat 180
aacataagaa tcctcttgac ggaaatatgc cttctcactc ctttacccct gaacagcctc 240
ccagaacgtt tacgacggca ccgacttccg caacgtctgc gccaaaaccg cgcccagcc 300
tctctgctgc gtggcccccg ttgtaagttg atgccccagc tcaagctcca gtctttggca 360
aacccattct gacacccaga ctgcaggccg gccaggctct tctgtgccag accgccgtcg 420
gtgcttga                                                       428


<210> 2
<211> 78
<212> DNA
<213> Artificial Sequence

```
<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 2
tcgggcacta cgtgccagta tagcaacgac tactactcgc aatgccttgt tccgcgtggc 60
tctagttctg gaaccgca                                               78


<210> 3
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 3
tcgtacggat cctcaagcac cgacggcggt                                  30


<210> 4
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 4
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagag caacggcaac 60
ggc                                                               63


<210> 5
<211> 2211
<212> DNA
<213> Trichoderma reesei

<220>
<221> promoter
<222> (1)..(2211)
<223> cbh1 promoter sequence

<400> 5
gaattctcac ggtgaatgta ggccttttgt agggtaggaa ttgtcactca agcaccccca 60
acctccatta cgcctccccc atagagttcc caatcagtga gtcatggcac tgttctcaaa 120
tagattgggg agaagttgac ttccgcccag agctgaaggt cgcacaaccg catgatatag 180
ggtcggcaac ggcaaaaaag cacgtggctc accgaaaagc aagatgtttg cgatctaaca 240
tccaggaacc tggatacatc catcatcacg cacgaccact ttgatctgct ggtaaactcg 300
tattcgccct aaaccgaagt gcgtggtaaa tctacacgtg gccccctttc ggtatactgc 360
gtgtgtcttc tctaggtgca ttctttcctt cctctagtgt tgaattgttt gtgttgggag 420
tccgagctgt aactacctct gaatctctgg agaatggtgg actaacgact accgtgcacc 480
tgcatcatgt atataatagt gatcctgaga aggggggttt ggagcaatgt gggactttga 540
tggtcatcaa acaaagaacg aagacgcctc ttttgcaaag ttttgtttcg ctacggtga 600
agaactggat acttgttgtg tcttctgtgt attttgtgg caacaagagg ccagagacaa 660
tctattcaaa caccaagctt gctctttttga gctacaagaa cctgtggggt atatatctag 720
agttgtgaag tcggtaatcc cgctgtatag taatacgagt cgcatctaaa tactccgaag 780
```

```
ctgctgcgaa cccggagaat cgagatgtgc tggaaagctt ctagcgagcg gctaaattag 840
catgaaaggc tatgagaaat tctggagacg gcttgttgaa tcatggcgtt ccattcttcg 900
acaagcaaag cgttccgtcg cagtagcagg cactcattcc cgaaaaaact cggagattcc 960
taagtagcga tggaaccgga ataatataat aggcaataca ttgagttgcc tcgacggttg 1020
caatgcaggg gtactgagct tggacataac tgttccgtac cccacctctt ctcaaccttt 1080
ggcgtttccc tgattcagcg tacccgtaca agtcgtaatc actattaacc cagactgacc 1140
ggacgtgttt tgcccttcat ttggagaaat aatgtcattg cgatgtgtaa tttgcctgct 1200
tgaccgactg gggctgttcg aagcccgaat gtaggattgt tatccgaact ctgctcgtag 1260
aggcatgttg tgaatctgtg tcgggcagga cacgcctcga aggttcacgg caagggaaac 1320
caccgatagc agtgtctagt agcaacctgt aaagccgcaa tgcagcatca ctggaaaata 1380
caaaccaatg ctaaaagta cataagttaa tgcctaaaga agtcatatac cagcggctaa 1440
taattgtaca atcaagtggc taaacgtacc gtaatttgcc aacgcgttgt ggggttgcag 1500
aagcaacggc aaagcccact tcccacgttt gtttcttcac tcagtccaat ctcagctggt 1560
gatcccccaa ttgggtcgct tgtttgttcc ggtgaagtga aagaagacag aggtaagaat 1620
gtctgactcg gagcgttttg catacaacca agggcagtga tggaagacag tgaaatgttg 1680
acattcaagg agtatttagc cagggatgct tgagtgtatc gtgtaaggag gtttgtctgc 1740
cgatacgacg aatactgtat agtcacttct gatgaagtgg tccatattga aatgtaagtc 1800
ggcactgaac aggcaaaaga ttgagttgaa actgcctaag atctcgggcc ctcgggcttc 1860
ggctttgggt gtacatgttt gtgctccggg caaatgcaaa gtgtggtagg atcgacacac 1920
tgctgccttt accaagcagc tgagggtatg tgataggcaa atgttcaggg gccactgcat 1980
ggtttcgaat agaaagagaa gcttagccaa gaacaatagc cgataaagat agcctcatta 2040
aacgaaatga gctagtaggc aaagtcagcg aatgtgtata tataaaggtt cgaggtccgt 2100
gcctccctca tgctctcccc atctactcat caactcagat cctccaggag acttgtacac 2160
catcttttga ggcacagaaa cccaatagtc aaccgcggac tgcgcatcat g        2211
```


\<210\> 6
\<211\> 1588
\<212\> DNA
\<213\> Trichoderma reesei

\<220\>
\<221\> misc_feature
\<223\> T. reesei egl1 cDNA

\<400\> 6
```
ccccctatc ttagtccttc ttgttgtccc aaaatggcgc cctcagttac actgccgttg 60
accacggcca tcctggccat tgcccggctc gtcgccgccc agcaaccggg taccagcacc 120
cccgaggtcc atcccaagtt gacaacctac aagtgtacaa agtccggggg gtgcgtggcc 180
caggacacct cggtggtcct tgactggaac taccgctgga tgcacgacgc aaactacaac 240
tcgtgcaccg tcaacggcgg cgtcaacacc acgctctgcc ctgacgaggc gacctgtggc 300
aagaactgct tcatcgaggg cgtcgactac gccgcctcgg cgtcacgac tcgggcagc 360
agcctcacca tgaaccagta catgcccagc agctctggcg gctacagcag cgtctctcct 420
cggctgtatc tcctggactc tgacggtgag tacgtgatgc tgaagctcaa cggccaggag 480
ctgagcttcg acgtcgacct ctctgctctg ccgtgtggag agaacggctc gctctacctg 540
tctcagatgg acgagaacgg gggcgccaac cagtataaca cggccggtgc caactacggg 600
agcggctact gcgatgctca gtgccccgtc cagacatgga ggaacggcac cctcaacact 660
agccaccagg gcttctgctg caacgagatg gatatcctgg agggcaactc gaggcgaat 720
gccttgaccc ctcactcttg cacggccacg gcctgcgact ctgccggttg cggcttcaac 780
ccctatggca gcggctacaa aagctactac ggccccggag ataccgttga cacctccaag 840
accttcacca tcatcaccca gttcaacacg gacaacggct cgccctcggg caaccttgtg 900
agcatcaccc gcaagtacca gcaaaacggc gtcgacatcc cagcgccca gcccggcggc 960
gacaccatct cgtcctgccc gtccgcctca gcctacggcg gcctcgccac catgggcaag 1020
gccctgagca gcggcatggt gctcgtgttc agcatttgga cgacaacag ccagtacatg 1080
aactggctcg acagcggcaa cgccggcccc tgcagcagca ccgagggcaa cccatccaac 1140
atcctggcca acaaccccaa cacgcacgtc gtcttctcca acatccgctg gggagacatt 1200
gggtctacta cgaactcgac tgcgccccca ccccgcctg cgtccagcac gacgttttcg 1260
```

```
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagac tcactggggg 1320
cagtgcggtg gcattgggta cagcgggtgc aagacgtgca cgtcgggcac tacgtgccag 1380
tatagcaacg actactactc gcaatgcctt tagagcgttg acttgcctct ggtctgtcca 1440
gacggggca cgatagaatg cgggcacgca gggagctcgt agacattggg cttaatatat 1500
aagacatgct atgttgtatc tacattagca aatgacaaac aaatgaaaaa gaacttatca 1560
agcaaaaaaa aaaaaaaaa aaaaaaaa                                   1588


<210> 7
<211> 745
<212> DNA
<213> Trichoderma reesei

<220>
<221> terminator
<222> (1)..(745)
<223> T. reesei cbh1 terminator

<400> 7
ggacctaccc agtctcacta cggccagtgc ggcggtattg gctacagcgg ccccacggtc 60
tgcgccagcg gcacaacttg ccaggtcctg aacccttact actctcagtg cctgtaaagc 120
tccgtgcgaa agcctgacgc accggtagat tcttggtgag cccgtatcat gacggcggcg 180
ggagctacat ggccccgggt gatttatttt ttttgtatct acttctgacc ctttcaaat 240
atacggtcaa ctcatctttc actggagatg cggcctgctt ggtattgcga tgttgtcagc 300
ttggcaaatt gtggctttcg aaaacacaaa acgattcctt agtagccatg cattttaaga 360
taacggaata gaagaaagag gaaattaaaa aaaaaaaaa aacaaacatc ccgttcataa 420
cccgtagaat cgccgctctt cgtgtatccc agtaccacgt caaaggtatt catgatcgtt 480
caatgttgat attgttccgc cagtatggct ccacccccat ctccgcgaat ctcctcttct 540
cgaacgcggt agtggctgct gccaattggt aatgaccata gggagacaaa cagcataata 600
gcaacagtgg aaattagtgg cgcaataatt gagaacacag tgagaccata gctggcggcc 660
tggaaagcac tgttggagac caacttgtcc gttgcgaggc caacttgcat tgctgtcaag 720
acgatgacaa cgtagccgag gaccc                                     745


<210> 8
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: annealed primer

<400> 8
taaccgcggt                                                       10


<210> 9
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: annealed primer

<400> 9
ctagaccgcg gttaat                                                16
```

```
<210> 10
<211> 1232
<212> DNA
<213> Trichoderma reesei

<220>
<221> promoter
<222> (1)..(1232)
<223> T. reesei gpd1 promotor

<400> 10
gtcgacacga tatacaggcg cggctgatga taatgatgat cgagcatgac ttgatgctgt 60
atgtgacaat attgactgcg aggaaccatc aggtgtgtat ggatggaatc attctgtaac 120
caccaaggtg catgcatcat aaggattctc ctcagctcac caacaacgaa cgatggccat 180
gttagtgaag gcaccgtgat ggcaagatag aaccactatt gcatctgcgc ttcccacgca 240
cagtacgtca agtaacgtca aagccgccct cccgtaacct cgccgttgt tgctcccccc 300
gattgcctca atcacatagt acctacctat gcattatggg cggcctcaac ccacccccc 360
agattgagag ctaccttaca tcaatatggc cagcacctct tcggcgatac atactcgcca 420
ccccagccgg cgcgattgtg tgtactaggt aggctcgtac tataccagca ggagaggtgc 480
tgcttggcaa tcgtgctcag ctgttaggtt gtacttgtat ggtacttgta aggtggtcat 540
gcagttgcta aggtacctag ggagggattc aacgagccct gcttccaatg tccatctgga 600
taggatggcg gctggcgggg ccgaagctgg gaactcgcca acagtcatat gtaatagctc 660
aagttgatga taccgttttg ccagattaga tgcgagaagc agcatgaatg tcgctcatcc 720
gatgccgcat caccgttgtg tcagaaacga ccaagctaag caactaaggt accttaccgt 780
ccactatctc aggtaaccag gtactaccag ctaccctacc tgccgtgcct acctgcttta 840
gtgttaatct ttccacctcc ctcctcaatc ttcttttccc tcctctcctc tttttttttt 900
cttcctcctc ttcttctcca taaccattcc taacaacatc gacattctct cctaatcacc 960
agcctcgcaa atcctcagtt tgtatgtacg tacgtactac aatcatcacc acgatcgtcc 1020
gcccgacgat gcggcttctg ttcgcctgcc cctcctctca ctcgtgccct tgacgagcta 1080
gcccccgccag gactctcctg cgtcaccaat ttttttccct atttacccct cctccctctc 1140
tccctctcgt ttcttcctaa caaacaacca ccaccaaaat ctctttggaa gctcacgact 1200
cacgcagctc aattcgcaga tacaaatcta ga                                1232


<210> 11
<211> 1129
<212> DNA
<213> Trichoderma reesei

<220>
<221> terminator
<222> (1)..(1129)
<223> T. reesei gpd1 terminator

<400> 11
ggatcccgag cattgtctat gaatgcaaac aaaaatagta aataaatagt aattctggcc 60
atgacgaata gagccaatct gctccacttg actatcttgt gactgtatcg tatgtcgaac 120
ccttgactgc ccattcaaac aattgtaaag gaatatagct acaagttatg tctcacgttt 180
gcgtgcgagc ccgtttgtac gttattttga gaaagcgttg ccatcacatg ctcacagtca 240
cttggcttac gatcatgttt gcgatcttcg gtaagaatac acagagtaac gattatctcc 300
atcgcttcta tgattaggta ctcagacaac acatgggaaa caagataacc atcgcatgca 360
aggtcgattc caatcatgat ctggactggg gtattccatc taagccatag taccctcgag 420
agaaggaatg gtaggacctc tcaggcgtcc accatctgtg ctgcaaatcc aagaaacccc 480
ccaaaagcac ctacctatct acctagagta actgcacgag aaaagaaaag gagcagaaga 540
agaatgatct caagaggccg tgaacgcaga aacacactcc tcccaacttt tcaagttttg 600
aacaaaaaaa gaaagatgag gactagaaga tggagtattt ccttcttaga gagctctcgg 660
```

```
tgaggtgacc tgtcagggtt taccgcaaac cgtcggtggt ·tctatccaat taatcaagtc 720
ccgcgcctcg cctcttctct cctgtccttt catagaatcc cgtctccttg ttgcttgatc 780
gaagcggggt tatcgacgcc accaaagatc ttgtcttggt gacttatcaa tcctttggtg 840
atcaaacagc ccccgagtga tcagatccgt aaaagaagaa gaagagtacg atttaaccag 900
accgaggaac aataaagcga gtaaataaca tcaaaataag agtctcgttg aaaattactt 960
gttcctcaat caatcccaac ccccctaaaa gcccttcccc ccatggtata tcccggcagt 1020
aggagagaga tatttccact accgctcacc accaagtgag gcttgccgag agaagaggat 1080
gaatcagaag tgacaacaac gggttgagca catgggatat cggcgcgcc         1129


<210> 12
<211> 5733
<212> DNA
<213> Aspergillus nidulans

<220>
<221> misc_feature
<223> (1-5733) Sequence of plasmid pAN52-1

<220>
<221> promoter
<222> (1)..(2129)
<223> A. nidulans gpdA promoter

<220>
<221> gene
<222> (2130)..(2304)
<223> A. nidulans gpdA gene

<220>
<221> terminator
<222> (2305)..(3071)
<223> A. nidulans trpC terminator

<220>
<221> misc_feature
<222> (3072)..(5726)
<223> pUC18 from SalI to EcoRI

<400> 12
caattccctt gtatctctac acacaggctc aaatcaataa gaagaacggt tcgtcttttt 60
cgtttatatc ttgcatcgtc ccaaagctat tggcgggata ttctgtttgc agttggctga 120
cttgaagtaa tctctgcaga tctttcgaca ctgaaatacg tcgagcctgc tccgcttgga 180
agcggcgagg agcctcgtcc tgtcacaact accaacatgg agtacgataa gggccagttc 240
cgccagctca ttaagagcca gttcatgggc gttggcatga tggccgtcat gcatctgtac 300
ttcaagtaca ccaacgctct tctgatccag tcgatcatcc gctgaaggcg ctttcgaatc 360
tggttaagat ccacgtcttc gggaagccag cgactggtga cctccagcgt ccctttaagg 420
ctgccaacag ctttctcagc cagggccagc ccaagaccga caaggcctcc ctccagaacg 480
ccgagaagaa ctggaggggt ggtgtcaagg aggagtaagc tccttattga agtcggagga 540
cggagcggtg tcaagaggat attcttcgac tctgtattat agataagatg atgaggaatt 600
ggaggtagca tagcttcatt tggatttgct ttccaggctg agactctagc ttggagcata 660
gagggtcctt tggctttcaa tattctcaag tatctcgagt ttgaacttat ccctgtgaa 720
ccttttattc accaatgagc attggaatga acatgaatct gaggactgca atcgccatga 780
ggtttttcgaa atacatccgg atgtcgaagg cttggggcac ctgcgttggt tgaatttaga 840
acgtggcact attgatcatc cgatagctct gcaaagggcg ttgcacaatg caagtcaaac 900
gttgctagca gttccaggtg gaatgttatg atgagcattg tattaaatca ggagatatag 960
catgatctct agttagctca ccacaaaagt cagacggcgt aaccaaaagt cacacaacac 1020
```

6

```
aagctgtaag gatttcggca cggctacgga agacggagaa gccaccttca gtggactcga 1080
gtaccattta attctatttg tgtttgatcg agacctaata cagcccctac aacgaccatc 1140
aaagtcgtat agctaccagt gaggaagtgg actcaaatcg acttcagcaa catctcctgg 1200
ataaacttta agcctaaact atacagaata agataggtgg agagcttata ccgagctccc 1260
aaatctgtcc agatcatggt tgaccggtgc ctggatcttc ctatagaatc atccttattc 1320
gttgacctag ctgattctgg agtgacccag agggtcatga cttgagccta aaatccgccg 1380
cctccaccat ttgtagaaaa atgtgacgaa ctcgtgagct ctgtacagtg accggtgact 1440
ctttctggca tgcggagaga cggacggacg cagagagaag ggctgagtaa taagccactg 1500
gccagacagc tctggcggct ctgaggtgca gtggatgatt attaatccgg gaccggccgc 1560
ccctccgccc cgaagtggaa aggctggtgt gcccctcgtt gaccaagaat ctattgcatc 1620
atcggagaat atggagcttc atcgaatcac cggcagtaag cgaaggagaa tgtgaagcca 1680
ggggtgtata gccgtcggcg aaatagcatg ccattaacct aggtacagaa gtccaattgc 1740
ttccgatctg gtaaaagatt cacgagatag taccttctcc gaagtaggta gagcgagtac 1800
ccggcgcgta agctccctaa ttggcccatc cggcatctgt agggcgtcca aatatcgtgc 1860
ctctcctgct ttgcccggtg tatgaaaccg gaaaggccgc tcaggagctg gccagcggcg 1920
cagaccggga acacaagctg gcagtcgacc catccggtgc tctgcactcg acctgctgag 1980
gtccctcagt ccctggtagg cagctttgcc ccgtctgtcc gcccggtgtg tcggcggggt 2040
tgacaaggtc gttgcgtcag tccaacattt gttgccatat tttcctgctc tccccaccag 2100
ctgctctttt cttttctctt tcttttccca tcttcagtat attcatcttc ccatccaaga 2160
acctttattt cccctaagta agtactttgc tacatccata ctccatcctt cccatccctt 2220
attcctttga acctttcagt tcgagctttc ccacttcatc gcagcttgac taacagctac 2280
cccgcttgag cagacatcac catggatcca cttaacgtta ctgaaatcat caaacagctt 2340
gacgaatctg gatataagat cgttggtgtc gatgtcagct ccggagttga dacaaatggt 2400
gttcaggatc tcgataagat acgttcattt gtccaagcag caaagagtgc cttctagtga 2460
tttaatagct ccatgtcaac aagaataaaa cgcgttttcg ggtttacctc ttccagatac 2520
agctcatctg caatgcatta atgcattgac tgcaacctag taacgccttn caggctccgg 2580
cgaagagaag aatagcttag cagagctatt ttcatttttcg ggagacgaga tcaagcagat 2640
caacggtcgt caagagacct acgagactga ggaatccgct cttggctcca cgcgactata 2700
tatttgtctc taattgtact ttgacatgct cctcttcttt actctgatag cttgactatg 2760
aaaattccgt caccagcncc tgggttcgca aagataattg catgtttctt ccttgaactc 2820
tcaagcctac aggacacaca ttcatcgtag gtataaacct cgaaatcant tcctactaag 2880
atggtataca atagtaacca tgcatggttg cctagtgaat gctccgtaac acccaatacg 2940
ccggccgaaa cttttttaca actctcctat gagtcgttta cccagaatgc acaggtacac 3000
ttgtttagag gtaatccttc tttctagaag tcctcgtgta ctgtgtaagc gcccactcca 3060
catctccact cgacctgcag gcatgcaagc ttggcactgg ccgtcgtttt acaacgtcgt 3120
gactgggaaa accctggcgt tacccaactt aatcgccttg cagcacatcc ccctttcgcc 3180
agctggcgta atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg 3240
aatggcgaat ggcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac 3300
cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga 3360
cacccgccaa cacccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac 3420
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg 3480
aaacgcgcga dacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata 3540
ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aaccccctatt 3600
tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 3660
atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 3720
attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 3780
gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 3840
agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 3900
aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 3960
cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 4020
cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 4080
actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 4140
cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 4200
ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 4260
ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 4320
gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 4380
gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat 4440
```

7

```
ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 4500
cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 4560
caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 4620
taggtgaaga tccttttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 4680
cactgagcgt cagacccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg 4740
cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 4800
gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 4860
aatactgtcc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 4920
cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 4980
tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg tcgggctga 5040
acggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 5100
ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 5160
ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 5220
tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg attttttgtga 5280
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 5340
ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg 5400
gataaccgta ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag 5460
cgcagcgagt cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc 5520
gcgcgttggc cgattcatta atgcagctgg cacgacaggt ttcccgactg gaaagcgggc 5580
agtgagcgca acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac 5640
tttatgcttc cggctcgtat gttgtgtgga attgtgagcg ataacaatt tcacacagga 5700
aacagctatg accatgatta cgaattgcgg ccg                               5733
```

```
<210> 13
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 13
gtcaaccgcg gactgcgcat catgaagttc ttcgccatc                        39
```

```
<210> 14
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 14
tctagcaagc ttggctctag ttctggaacc gcaccaggcg gcagcaacgg caacggcaat 60
gtttgc                                                            66
```

```
<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer
```

8

<400> 15
tcgtacaagc tttcaagcac cgacggcggt                    30


<210> 16
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 16
tctagctcta gaagcaacgg caacggcaat gtt               33


<210> 17
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 17
tgctagtcga cctgctagca gcaccgacgg cggtctg           37


<210> 18
<211> 4614
<212> DNA
<213> Saccharomyces cerevisiae

<220>
<223> S. cerevisiae FLO1 coding sequence

<400> 18
atgacaatgc ctcatcgcta tatgtttttg gcagtcttta cacttctggc actaactagt 60
gtggcctcag gagccacaga ggcgtgctta ccagcaggcc agaggaaaag tgggatgaat 120
ataaattttt accagtattc attgaaaagat tcctccacat attcgaatgc agcatatatg 180
gcttatggat atgcctcaaa aaccaaacta ggttctgtcg gaggacaaac tgatatctcg 240
attgattata atattccctg tgttagttca tcaggcacat ttccttgtcc tcaagaagat 300
tcctatggaa actggggatg caaaggaatg ggtgcttgtt ctaatagtca aggaattgca 360
tactggagta ctgatttatt tggtttctat actaccccaa caaacgtaac cctagaaatg 420
acaggttatt ttttaccacc acagacgggt tcttacacat tcaagtttgc tacagttgac 480
gactctgcaa ttctatcagt aggtggtgca accgcgttca actgttgtgc tcaacagcaa 540
ccgccgatca catcaacgaa ctttaccatt gacggtatca agccatgggg tggaagtttg 600
ccacctaata tcgaaggaac cgtctatatg tacgctggct actattatcc aatgaaggtt 660
gtttactcga acgctgtttc ttggggtaca cttccaatta gtgtgacact tccagatggt 720
accactgtaa gtgatgactt cgaagggtac gtctattcct tgacgatga cctaagtcaa 780
tctaactgta ctgtccctga cccttcaaat tatgctgtca gtaccactac aactacaacg 840
gaaccatgga ccggtacttt cacttctaca tctactgaaa tgaccaccgt caccggtacc 900
aacggcgttc caactgacga aaccgtcatt gtcatcagaa ctccaacaac tgctagcacc 960
atcataacta caactgagcc atggaacagc actttttacct ctacttctac cgaattgacc 1020
acagtcactg gcaccaatgg tgtacgaact gacgaaacca tcattgtaat cagaacacca 1080
acaacagcca ctactgccat aactacaact gagccatgga acagcacttt tacctctact 1140
tctaccgaat tgaccacagt caccggtacc aatggtttgc aactgatga gaccatcatt 1200


9

```
gtcatcagaa caccaacaac agccactact gccatgacta caactcagcc atggaacgac 1260
actttacct ctacatccac tgaaatgacc accgtcaccg gtaccaacgg tttgccaact 1320
gatgaaacca tcattgtcat cagaacacca acaacagcca ctactgctat gactacaact 1380
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccaccgt caccggtacc 1440
aacggtttgc caactgatga aaccatcatt gtcatcagaa caccaacaac agccactact 1500
gccatgacta caactcagcc atggaacgac actttacct ctacatccac tgaaatgacc 1560
accgtcaccg gtaccaatgg tttgccaact gatgagacca tcattgtcat cagaacacca 1620
acaacagcca ctactgccat gactacaact cagccatgga acgacacttt tacctctaca 1680
tccactgaaa tgaccaccgt caccggtacc aacggtttgc caactgatga aaccatcatt 1740
gtcatcagaa caccaacaac agccactact gccataacta caactgagcc atggaacagc 1800
actttacct ctacttctac cgaattgacc acagtcaccg gtaccaatgg tttgccaact 1860
gatgagacca tcattgtcat cagaacacca acaacagcca ctactgccat gactacaact 1920
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccaccgt caccggtacc 1980
aacggtttgc caactgatga aaccatcatt gtcatcagaa caccaacaac agccactact 2040
gccatgacta caactcagcc atggaacgac actttacct ctacatccac tgaaatgacc 2100
accgtcaccg gtaccaacgg tttgccaact gatgagacca tcattgtcat cagaacacca 2160
acaacagcca ctactgccat gactacaact cagccatgga acgacacttt tacctctaca 2220
tccactgaaa tgaccaccgt caccggtacc aacggcgttc caactgacga aaccgtcatt 2280
gtcatcagaa ctccaactag tgaaggtcta atcagcacca ccactgaacc atggactggt 2340
actttcacct ctacatccac tgagatgacc accgtcaccg gtactaacgg tcaaccaact 2400
gacgaaaccg tgattgttat cagaactcca accagtgaag gtttggttac aaccaccact 2460
gaaccatgga ctggtacttt tacttctaca tctactgaaa tgaccaccat tactggaacc 2520
aacggcgttc caactgacga aaccgtcatt gtcatcagaa ctccaaccag tgaaggtcta 2580
atcagcacca ccactgaacc atggactggt actttactt ctacatctac tgaaatgacc 2640
accattactg gaaccaatgg tcaaccaact gacgaaaccg ttattgttat cagaactcca 2700
actagtgaag gtctaatcag caccaccact gaaccatgga ctggtacttt cacttctaca 2760
tctactgaaa tgaccaccgt caccggtacc aacggcgttc caactgacga aaccgtcatt 2820
gtcatcagaa ctccaaccag tgaaggtcta atcagcacca ccactgaacc atggactggc 2880
actttcactt cgacttccac tgaggttacc accatcactg gaaccaacgg tcaaccaact 2940
gacgaaactg tgattgttat cagaactcca accagtgaag gtctaatcag caccaccact 3000
gaaccatgga ctggtacttt cacttctaca tctgctgaaa tgaccaccgt caccggtact 3060
aacggtcaac caactgacga aaccgtgatt gttatcagaa ctccaaccag tgaaggtttg 3120
gttacaacca ccactgaacc atggactggt actttactt cgacttccac tgaaatgtct 3180
actgtcactg gaaccaatgg cttgccaact gatgaaactg tcattgttgt caaaactcca 3240
actactgcca tctcatccag tttgtcatca tcatcttcag gacaaatcac cagctctatc 3300
acgtcttcgc gtccaattat tacccccattc tatcctagca atggaacttc tgtgatttct 3360
tcctcagtaa tttcttcctc agtcacttct tctctattca cttcttctcc agtcatttct 3420
tcctcagtca tttcttcttc tacaacaacc tccacttcta tattttctga atcatctaaa 3480
tcatccgtca ttccaaccag tagttccacc tctggttctt ctgagagcga aacgagttca 3540
gctggttctg tctcttcttc ctcttttatc tcttctgaat catcaaaatc tcctacatat 3600
tcttcttcat cattaccact tgttaccagt gcgacaacaa gccaggaaac tgcttcttca 3660
ttaccacctg ctaccactac aaaaacgagc gaacaaacca ctttggttac cgtgacatcc 3720
tgcgagtctc atgtgtgcac tgaatccatc tcccctgcga ttgtttccac agctactgtt 3780
actgttagcg gcgtcacaac agagtatacc acatggtgcc ctatttctac tacagagaca 3840
acaaagcaaa ccaaagggac aacagagcaa accacagaaa caacaaaaca aaccacggta 3900
gttacaattt cttcttgtga atctgacgta tgctctaaga ctgcttctcc agccattgta 3960
tctacaagca ctgctactat taacggcgtt actacagaat acacaacatg gtgtcctatt 4020
tccaccacag aatcgaggca acaaacaacg ctagttactg ttacttcctg cgaatctggt 4080
gtgtgttccg aaactgcttc acctgccatt gtttcgacgg ccacggctac tgtgaatgat 4140
gttgttacgg tctatcctac atggaggcca cagactgcga tgaagagtc tgtcagctct 4200
aaaatgaaca gtgctaccgg tgagacaaca accaatactt tagctgctga aacgactacc 4260
aatactgtag ctgctgagac gattaccaat actggagctg ctgagacgaa aacagtagtc 4320
acctcttcgc tttcaagatc taatcacgct gaaacacaga cggcttccgc gaccgatgtg 4380
attggtcaca gcagtagtgt tgtttctgta tccgaaactg gcaacaccaa gagtctaaca 4440
agttccgggt tgagtactat gtcgcaacag cctcgtagca caccagcaag cagcatggta 4500
ggatatagta cagcttcttt agaaatttca acgtatgctg gcagtgccaa cagcttactg 4560
gccggtagtg gtttaagtgt cttcattgcg tccttattgc tggcaattat ttaa       4614
```

10

<210> 19
<211> 5857
<212> DNA
<213> Saccharomyces cerevisiae and E. coli

<220>
<221> promoter
<222> (1)..(452)
<223> S. cerevisiae GAL1 promoter

<220>
<221> misc_feature
<223> (476-495) E. coli T7 promoter/priming site

<220>
<221> misc_feature
<223> (502-601) E. coli multiple cloning site

<220>
<221> misc_feature
<223> (609-857) S. cerevisiae CYC1 transcription terminator

<220>
<221> misc_feature
<223> (1039-1712) E. coli pMB1 (pUC-derived) origin

<220>
<221> gene
<222> (1857)..(2717)
<223> E. coli ampicillin resistance gene

<220>
<221> gene
<222> (2735)..(3842)
<223> S. cerevisiae URA3 gene

<220>
<221> misc_feature
<223> (3846-5317) S. cerevisiae 2 micron origin

<220>
<221> misc_feature
<223> (5385-5840) E. coli f1 origin

<220>
<221> misc_feature
<223> (1-5857) Sequence of pYES2

<400> 19
acggattaga agccgccgag cgggtgacag ccctccgaag gaagactctc ctccgtgcgt 60
cctcgtcctc accggtcgcg ttcctgaaac gcagatgtgc ctcgcgccgc actgctccga 120
acaataaaga ttctacaata ctagctttta tggttatgaa gaggaaaaat tggcagtaac 180
ctggccccac aaaccttcaa atgaacgaat caaattaaca accataggat gataatgcga 240
ttagtttttt agccttattt ctggggtaat taatcagcga agcgatgatt tttgatctat 300
taacagatat ataaatgcaa aaactgcatt aaccactttta actaatactt tcaacatttt 360

11

```
cggtttgtat tacttcttat tcaaatgtaa taaaagtatc aacaaaaaat tgttaatata 420
cctctatact ttaacgtcaa ggagaaaaaa ccccggatcg gactactagc agctgtaata 480
cgactcacta tagggaatat taagcttggt accgagctcg gatccactag taacggccgc 540
cagtgtgctg gaattctgca gatatccatc acactggcgg ccgctcgagc atgcatctag 600
agggccgcat catgtaatta gttatgtcac gcttacattc acgccctccc cccacatccg 660
ctctaaccga aaaggaagga gttagacaac ctgaagtcta ggtccctatt tattttttta 720
tagttatgtt agtattaaga acgttatTta tatttcaaat ttttcttttt tttctgtaca 780
gacgcgtgta cgcatgtaac attatactga aaaccttgct tgagaaggtt ttgggacgct 840
cgaaggcttt aatttgcggc cctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 900
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 960
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 1020
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa agcccaggaa ccgtaaaaag 1080
gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga 1140
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 1200
ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 1260
tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 1320
gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 1380
tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 1440
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 1500
ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 1560
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 1620
accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 1680
tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 1740
cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat 1800
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 1860
caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 1920
gcctgactcc ccgtcgtgta gataactacg atacgggagc gcttaccatc tggccccagt 1980
gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 2040
ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc cattcagtct 2100
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt 2160
gttggcattg ctacaggcat cgtggtgtca ctctcgtcgt ttggtatggc ttcattcagc 2220
tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt 2280
agctccttcg gtcctccgat cgttgtcaga gtaagttgg ccgcagtgtt atcactcatg 2340
gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg 2400
actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct 2460
tgcccggcgt caatacggga taatagtgta tcacatagca gaactttaaa agtgctcatc 2520
attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt 2580
tcgatgtaac ccactcgtgc acccaactga tcttcagcat ctttttactt caccagcgtt 2640
tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg 2700
aaatgttgaa tactcatact cttcctttt caatgggtaa taactgatat aattaaattg 2760
aagctctaat ttgtgagttt agtatacatg catttactta taatacagtt ttttagtttt 2820
gctggccgca tcttctcaaa tatgcttccc agcctgcttt tctgtaacgt tcaccctcta 2880
ccttagcatc ccttcccttt gcaaatagtc ctcttccaac aataataatg tcagatcctg 2940
tagagaccac atcatccacg gttctatact gttgacccaa tgcgtctccc ttgtcatcta 3000
aacccacacc gggtgtcata atcaaccaat cgtaaccttc atctcttcca cccatgtctc 3060
tttgagcaat aaagccgata acaaaatctt gtcgctctt cgcaatgtca acagtaccct 3120
tagtatattc tccagtagat agggagccct tgcatgacaa ttctgctaac atcaaaaggc 3180
ctctaggttc ctttgttact tcttctgccg cctgcttcaa accgctaaca atacctgggc 3240
ccaccacacc gtgtgcattc gtaatgtctg cccattctgc tattctgtat acacccgcag 3300
agtactgcaa tttgactgta ttaccaatgt cagcaaattt tctgtcttcg aagagtaaaa 3360
aattgtactt ggcggataat gcctttagcg gcttaactgt gccctccatg gaaaaatcag 3420
tcaagatatc cacatgtgtt tttagtaaac aaattttggg acctaatgct tcaactaact 3480
ccagtaattc cttggtggta cgaacatcca atgaagcaca caagtttgtt tgcttttcgt 3540
gcatgatatt aaatagcttg gcagcaacag gactaggatg agtagcagca cgttccttat 3600
atgtagcttt cgacatgatt tatcttcgtt tcctgcaggt ttttgttctg tgcagttggg 3660
ttaagaatac tgggcaattt catgtttctt caacactaca tatgcgtata tataccaatc 3720
taagtctgtg ctccttcctt cgttcttcct tctgttcgga gattaccgaa tcaaaaaaat 3780
```

12

```
ttcaaagaaa ccgaaatcaa aaaaaagaat aaaaaaaaaa tgatgaattg aattgaaaag 3840
ctagcttatc gatgataagc tgtcaaagat gagaattaat tccacggact atagactata 3900
ctagatactc cgtctactgt acgatacact tccgctcagg tccttgtcct ttaacgaggc 3960
cttaccactc ttttgttact ctattgatcc agctcagcaa aggcagtgtg atctaagatt 4020
ctatcttcgc gatgtagtaa aactagctag accgagaaag agactagaaa tgcaaaaggc 4080
acttctacaa tggctgccat cattattatc cgatgtgacg ctgcagcttc tcaatgatat 4140
tcgaatacgc tttgaggaga tacagcctaa tatccgacaa actgttttac agatttacga 4200
tcgtacttgt tacccatcat tgaattttga acatccgaac ctgggagttt tccctgaaac 4260
agatagtata tttgaacctg tataataata tatagtctag cgctttacgg aagacaatgt 4320
atgtatttcg gttcctggag aaactattgc atctattgca taggtaatct tgcacgtcgc 4380
atccccggtt cattttctgc gtttccatct tgcacttcaa tagcatatct ttgttaacga 4440
agcatctgtg cttcattttg tagaacaaaa atgcaacgcg agagcgctaa ttttttcaaac 4500
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tattttacca 4560
acgaagaatc tgtgcttcat ttttgtaaaa caaaaatgca acgcgacgag agcgctaatt 4620
tttcaaacaa agaatctgag ctgcattttt acagaacaga aatgcaacgc gagagcgcta 4680
ttttaccaac aaagaatcta tacttctttt ttgttctaca aaaatgcatc ccgagagcgc 4740
tattttttcta acaaagcatc ttagattact ttttttctcc tttgtgcgct ctataatgca 4800
gtctcttgat aacttttttgc actgtaggtc cgttaaggtt agaagaaggc tactttggtg 4860
tctattttct cttccataaa aaaagcctga ctccacttcc cgcgtttact gattactagc 4920
gaagctgcgg gtgcattttt tcaagataaa ggcatccccg attatattct ataccgatgt 4980
ggattgcgca tactttgtga acagaaagtg atagcgttga tgattcttca ttggtcagaa 5040
aattatgaac ggtttcttct attttgtctc tatatactac gtataggaaa tgtttacatt 5100
ttcgtattgt tttcgattca ctctatgaat agttcttact acaattttttt tgtctaaaga 5160
gtaatactag agataaacat aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag 5220
cgaaaggtgg atgggtaggt tatataggga tatagcacag agatatatag caaagagata 5280
cttttgagca atgtttgtgg aagcggtatt cgcaatggga agctccaccc cggttgataa 5340
tcagaaaagc cccaaaaaca ggaagattgt ataagcaaat atttaaattg taaacgttaa 5400
tattttgtta aaattcgcgt taaattttttg ttaaatcagc tcattttttta acgaatagcc 5460
cgaaatcggc aaaatccctt ataaatcaaa agaatagacc gagatagggt tgagtgttgt 5520
tccagtttcc aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa 5580
aagggtctat caggcgatg gcccactacg tgaaccatca ccctaatcaa gttttttggg 5640
gtcgaggtgc cgtaaagcag taaatcggaa gggtaaacgg atgcccccat ttagagcttg 5700
acggggaaag ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcggggggc 5760
tagggcggtg gaagtgtag gggtcacgct gggcgtaacc accacacccg ccgcgcttaa 5820
tgggcgcta cagggcgcgt ggggatgatc cactagt              5857
```

<210> 20
<211> 403
<212> DNA
<213> Trichoderma reesei

<220>
<221> misc_feature
<223> (1-403) T. reesei hfb2 coding sequence

<220>
<221> intron
<222> (131)..(200)

<220>
<221> intron
<222> (287)..(358)

<400> 20
```
atgcagttct tcgccgtcgc cctcttcgcc accagcgccc tggctgctgt ctgccctacc 60
ggcctcttct ccaaccctct gtgctgtgcc accaacgtcc tcgacctcat tggcgttgac 120
```

13

```
tgcaagaccc gtatgttgaa ttccaatctc tgggcatcct gacattggac gatacagttg 180
acttacacga tgctttacag ctaccatcgc cgtcgacact ggcgccatct tccaggctca 240
ctgtgccagc aagggctcca agcctctttg ctgcgttgct cccgtggtaa gtagtgctcg 300
caatggcaaa gaagtaaaaa gacatttggg cctgggatcg ctaactcttg atatcaaggc 360
cgaccaggct ctcctgtgcc agaaggccat cggcaccttc taa             403


<210> 21
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 21
cggaggagct cgacgacttc gagcagcccg agctgcacgc aggctgtctg ccctaccgg  59


<210> 22
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 22
tcattggatc cttagaaggt gccgatggc                                  29


<210> 23
<211> 679
<212> DNA
<213> Schizophyllum commune

<220>
<221> misc_feature
<223> (1-679) SC3 coding sequence

<220>
<221> misc_feature
<223> (1-92) 1st cDNA

<220>
<221> misc_feature
<223> (146-183) 2nd cDNA

<220>
<221> misc_feature
<223> (240-317) 3rd cDNA

<220>
<221> misc_feature
<223> (374-469) 4th cDNA

<220>
```

<221> misc_feature
<223> (524-586) 5th cDNA

<220>
<221> misc_feature
<223> (635-679) 6th cDNA

<400> 23
atgttcgccc gtctccccgt cgtgttcctc tacgccttcg tcgcgttcgg cgccctcgtc 60
gctgccctcc caggtggcca cccgggcacg acgtacgtcg acctctcacc gtcctctaat 120
gtcttgctga tgaagccccg tatagcacgc cgccggttac gacgacggtg acggtgacca 180
cggtgagtag ctttctcgcc gtcgacgact cgaacgcatt ggctaatttt tgctcatagc 240
cgccctcgac gacgaccatc gccgccggtg gcacgtgtac tacggggtcg ctctcttgct 300
gcaaccaggt tcaatcggta cgtacatcaa agcggcacga ccaggcatct cagctgacgg 360
ccacatcgta caggcgagca gcagccctgt taccgccctc ctcggcctgc tcggcattgt 420
cctcagcgac ctcaacgttc tcgttggcat cagctgctct cccctcactg tgagatcttt 480
ttgttcactg tcccaattac tgcgcactga cagactttgc caggtcatcg gtgtcggagg 540
cagcggctgt tcggcgcaga ccgtctgctg cgaaaacacc caattcgtat gtatactttc 600
catgcgtgtc cctttctccg ctaatcatct gtagaacggg ctgatcaaca tcggttgcac 660
ccccatcaac atcctctga                                               679


<210> 24
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 24
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcaggg tggccacccg 60
ggc                                                                63


<210> 25
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 25
tcgtacggat cctcagagga tgttgatggg                                   30


<210> 26
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 26
ggaattccgc ggactgcgca tcatgaagtt cttcgccatc gcc                    43


15

```
<210> 27
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 27
tgaattccat atgttaggta ccaccggggc ccatgccggt agaagtagaa gccccgggag  60
caccgacggc ggtctggcac                                               80


<210> 28
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 28
tgaattcggt acccaggctt gctcaagcgt c                                  31


<210> 29
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 29
tgaattccat atgtcacagg cactgagagt agta                               34


<210> 30
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 30
gaattcggta ccctcgtccc tcgcggtccc gccgaagtga acctggtg                48


<210> 31
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
```

<223> Description of Artificial Sequence: PCR 3' primer

<400> 31
tgaattccat atgctaaccc cgtttcatct ccag                               34


<210> 32
<211> 918
<212> DNA
<213> Trichoderma reesei

<220>
<221> terminator
<222> (1)..(918)
<223> T. reesei hfb1 terminator

<400> 32
gatgcccgcc cggggtcaag gtgtgcccgt gagaaagccc acaaagtgtt gatgaggacc 60
atttccggta ctgggaaagt tggctccacg tgtttgggca ggtttgggca agttgtgtag 120
atattccatt cgtacgccat tcttattctc caatatttca gtacactttt cttcataaat 180
caaaaagact gctattctct ttgtgacatg ccggaaggga acaattgctc ttggtctctg 240
ttatttgcaa gtaggagtgg gagattcgcc ttagagaaag tagagaagct gtgcttgacc 300
gtggtgtgac tcgacgagga tggactgaga gtgttaggat taggtcgaac gttgaagtgt 360
atacaggatc gtctggcaac ccacggatcc tatgacttga tgcaatggtg aagatgaatg 420
acagtgtaag aggaaaagga aatgtccgcc ttcagctgat atccacgcca atgatacagc 480
gatatacctc caatatctgt gggaacgaga catgacatat ttgtgggaac aacttcaaac 540
agcgagccaa gacctcaata tgcacatcca aagccaaaca ttggcaagac gagagacagt 600
cacattgtcg tcgaaagatg gcatcgtacc caaatcatca gctctcatta tcgcctaaac 660
cacagattgt ttgccgtccc ccaactccaa aacgttacta caaaagacat gggcgaatgc 720
aaagacctga aagcaaaccc tttttgcgac tcaattccct cctttgtcct cggaatgatg 780
atccttcacc aagtaaaaga aaaagaagat tgagataata catgaaaagc acaacggaaa 840
cgaaagaacc aggaaaagaa taaatctatc acgcaccttg tccccacact aaaagcaaca 900
gggggggtaa aatgaaat                                                 918


<210> 33
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 33
gacctcgatg cccgcccggg gtcaag                                        26


<210> 34
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 34
gtcgacattt cattttaccc ccctcg                                        26

<210> 35
<211> 1190
<212> DNA
<213> Trichoderma reesei

<220>
<221> promoter
<222> (1)..(1190)
<223> T. reesei hfb2 promoter

<400> 35
ctcgagcagc tgaagcttgc atgcctgcat cctttgtgag cgactgcatc cattttgcac 60
acactgccgt cgacgtctct cttccgacct tggccagctg gacaagcaac acaccaatga 120
cgctttgtat tattagagta tatgcaagtc tcaggactat cgactcaact ctacccaccg 180
aggacgatcg cggcacgata cgccctcgtt ctcattggcc caagcagacc aactgcccct 240
ggagcaagat tcagcccaag ggagatggac ggcagggcac gccaggcccc caccaccaag 300
ccactccctt tggccaaatc agcttgcatg tcaagagaca tcgagctgtg ccttgaaatt 360
actaacaacc agggatggga aacgaagcct gcttttggaa agacaacaat gagagagaga 420
gagagaggga gagagacaat gagtgccaca aacctggtag tgctccgcca atgcgtctga 480
aatgtcacat ccgagtcttg gggcctctgt gagaatgtcc agagtaatac gtgttttgcg 540
aatagtcctc tttcttgagg actggatacc tacgataccc ttttgagtt gatgcggtgc 600
tttcgaagta ttatctggag gatagaagac gtctaggtaa ctacacaaaa ggcctatact 660
ttggggagta gcccaacgaa aggtaactcc tacggcctct tagagccgtc atagatccta 720
cagcctcttg gagccgtcat agatcacatc tgtgtagacc gacattctat gaataatcat 780
ctcatcatgg ccacatacta ctacatacgt gtctctgcct acctgacatg tagcagtggc 840
caagacacca aggccccagc atcaagcctc cctacctatc ccttccattg tacagcggca 900
gagagattgc gatgagccct ctccctacct acagacggct gacaatgtcc gtataccacc 960
agccaacgtg atgaaaacaa ggacatgagg aacagcctgc gagagctgga agatgaagag 1020
ggccagaaaa aaaagtataa agaagacctc gattcccgcc atccaacaat cttttccatc 1080
ctcatcagca cactcatcta caaccatcac cacattcact caactcctct ttctcaactc 1140
tccaaacaca aacattcttt gttgaatacc aaccatcacc acctttcaag 1190


<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 5' primer

<400> 36
aagcttgcat gcctgcatcc                                                    20


<210> 37
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR 3' primer

<400> 37
ccatggtgaa aggtggtgat ggttgg                                             26


18

```
<210> 38
<211> 13
<212> PRT
<213> Trichoderma reesei

<220>
<221> misc_feature
<223> vild type T. reesei EGI peptide linker

<400> 38
Val Pro Arg Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly
  1               5                   10



<210> 39
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: modified CBHII linker

<400> 39
Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly
  1               5                   10



<210> 40
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Met/Thrombin
      linker

<400> 40
Pro Gly Arg Pro Val Leu Thr Gly Pro Gly Met Gly Thr Ser Thr Ser
  1               5                   10                      15

Ala Gly Pro



<210> 41
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Met-containing linker

<400> 41
```

Pro Gly Ala Ser Thr Ser Thr Gly Met Gly Pro Gly Gly
 1               5                   10


<210> 42
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: linker containing the thrombin
 cleavage site

<400> 42
Gly Thr Leu Val Pro Arg Gly Pro Ala Gly Val Asn Leu Val
 1               5                   10


<210>  43
<211>  16
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  Description of Artificial Sequence: synthetic oligonucleotide
       NheIBgIIINheI of the pTNS15 plasmid

<400>  43
gctagagatc tctagc                                                    16


<210>  44
<211>  9
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Description of Artificial Sequence: synthetic peptide AocIXbaIAocI of
       the pTNS15 plasmid

<400>  44

Ala Ser Gly Ala Ser Arg Ala Ser Gly
1               5


<210>  45
<211>  27
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  Description of Artificial Sequence: synthetic oligonucleotide
       AocIXbaIAocI of the pTNS15 plasmid

<400>  45

20

gcctcaggag cctctagagc ttcagga                                              27


<210>   46
<211>   20
<212>   PRT
<213>   Trichoderma reesei

<400>   46

Ala Asn Ala Phe Cys Pro Glu Gly Leu Leu Tyr Thr Asn Pro Leu Cys
1                   5                   10                  15


Cys Asp Leu Leu
            20